

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 557

April, 1980

**AN INFORMATION PROCESSING APPROACH
TO UNDERSTANDING THE VISUAL CORTEX**

Francis H. C. Crick, David C. Marr and Tomaso Poggio

Abstract. An outline description is given of the experimental work on the visual acuity and hyperacuity of human beings. The very high resolution achieved in hyperacuity corresponds to a fraction of the spacing between adjacent cones in the fovea. We briefly outline a computational theory of early vision, according to which (a) retinal image is filtered through a set of approximately bandpass, spatial filters and (b) zero-crossings may contain sufficient information for much of the subsequent processing. Consideration of the optimum filter lead to one which is equivalent to a cell with a particular center-surround type of response. An "edge" in the visual field then corresponds to a line of zero-crossings in the filtered image. The mathematics of sampling and of Logan's zero-crossing theorem are briefly explained.

In this framework we suggest, similarly to Barlow (1979), that the fine grid of small cells in layer $4C\beta$ of the striate cortex perform an approximate reconstruction of the filtered image with the goal of representing the position of zero-crossings with a very high accuracy (around a few seconds of arc). How this might be achieved is discussed in outline. Finally it is mentioned that this picture is probably too static, as hyperacuity can be achieved with moving targets.

To make for easier comprehension much of the mathematics is set out in an appendix.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by National Science Foundation Grant MCS77-07569.

1. Introduction

That we see the world as well as we do is something of a miracle. What seems so direct and effortless turns out, on close consideration, to involve many rapid and complex processes the details of which we are only beginning to glimpse. In a series of papers Marr and his collaborators have outlined a general strategy for approaching some of these problems (Marr, 1976, 1977; Marr and Nishihara, 1978b; Marr and Poggio, 1976; Ullman, 1979). Two useful introductions are the general account by Marr and Nishihara (1978a) and a sympathetic review in *Nature* by Sutherland (1979). Other papers by Marr and his coworkers have made detailed suggestions about the computations the brain carries out in special cases (Marr and Poggio, 1976; Marr 1977; Ullman, 1979). In particular a recent paper by Marr and Poggio (1977, 1979) has proposed an algorithm for human stereopsis which has been successfully implemented in a computer program (Grimson and Marr, 1979). Computational theories of this kind based on old and new results of information theory can establish what needs to be computed and how, while psychophysical experiments can tell us, for instance, the precision of the computation. Additional constraints are determined by the biophysics of nerve cells and their recorded physiological properties as well as by the anatomical and physiological diagrams of the circuitry.

In this chapter we shall deal only with the very early stages in the main visual pathway, specifically the retina, the lateral geniculate nucleus (LGN) and the striate cortex (also called area 17 or V1). The visual system is attractive not only because it can be supplied with a well-controlled and detailed input (unlike the cerebellum) but also because it has a fairly simple and direct path from the sensory receptors to the cortex (unlike the auditory system). Moreover a large amount of detailed experimental work has been done on it, using a variety of methods, including such different approaches as neurophysiology, neuroanatomy and psychophysics.

To our regret we are not yet able to make detailed and explicit suggestions as to what all the neurons in this region are doing. Instead we shall try to explain the general way in which we are approaching these problems from a computational point of view. In particular we shall apply it to the phenomenon of hyperacuity since, as Barlow (1979) has pointed out, this raises special problems, both static and dynamic.

2. Acuity and Hyperacuity

The main experiments (which we summarize here only very briefly) have been carried out on human beings, especially in recent years, by Westheimer and his

collaborators (Westheimer, 1976, 1977; Westheimer and McKee, 1975, 1977a, b, 1978). If two points of light lie side by side they can be seen to be double. If they are put closer together they may appear to us as a single point. It is found that, with practice, an angular separation of about 1' of arc can be distinguished with 75% success (see for instance Westheimer, 1977). This is the classical test for two-point acuity. (See Figure 1(a).)

Much closer angular intervals can be detected in special situations. A typical example of this is the acuity found in reading a vernier. The objects used need not be straight lines. A pattern of three points also gives good results (see Figure 1b), the task being to say whether the middle point lies to the right or to the left of the imaginary line connecting the outer two points (Beck and Schwartz, 1978). In such tasks 75% success can be obtained, with practice, using an angular misalignment of only 2" to 5" of arc. That is, of a few seconds of arc rather than the minute of arc found for two-point acuity. Acuity of this type is often called hyperacuity, though 'positional accuracy' might be a better name. To achieve it the three points should not be too far apart — a few minutes of arc separation gives the best performance.

What is at the root of this large difference between hyperacuity and simple two-point acuity? For such high resolution the points must be adjacent but they must not overlap. This has been clearly pointed out by Westheimer and McKee (1977a).

Hyperacuity, that is positional accuracy, is found in a variety of situations. In particular the acuity observed in stereopsis is hyperacuity (Westheimer and McKee, 1978). This and other data (possibly Julesz, 1971, figs. 3.6.1 — 3.6.3) show that hyperacuity can be binocular. For this reason it is likely to be implemented no earlier in the visual pathway than the striate cortex, since this is the first place where the inputs from the two separate eyes interact strongly. The image need not be prolonged in time. A flash of 1.5 msec. is quite adequate and in fact the different parts of the pattern need not be flashed simultaneously provided they are not separated in time by more than 20 msec. (Westheimer and McKee, 1977a). Furthermore, the random line stereogram of Julesz (1971, Figure 3.6.1) would seem to imply that vernier acuity is not restricted to forced choice tests.

Most remarkable of all, as Barlow (1979) has emphasized in a recent note, performance in a hyperacuity task is not appreciably degraded even if the target is moving at rates up to 2° to 4° per second (Westheimer and McKee, 1975). This is not due to eye movements. Westheimer has used the technique of presenting the signal for only 200 msec., with the direction of motion randomized. This is too short a time for eye movements to be initiated correctly.

The astonishing nature of this performance (which even without practice is fairly striking) can be seen when the properties of the retina are considered. The spacing of the receptors, the cones, even in the fovea where they are closest, is about $25''$. In addition the optical spread of the system has a half-width of about $30''$. The spacing of the retinal ganglion cells, which connect the eye to the brain, is no finer. How then can we achieve such a high performance with such a blunt optical instrument?



Figure 1. (a) Pattern configuration for two-point acuity tests. (b) One of several patterns yielding hyperacuity. Lateral displacements of the middle point to the right or to the left of the imaginary lines connecting the outer 2 points can be detected down to about $5''$ accuracy. A typical separation for the dots is $10'$.

3. Theory

To approach this problem we need to use a computational approach. The underlying idea is that the nervous system is a very complex information processing machine. While we cannot always say exactly how visual information is handled we can be sure that once information is lost in a pathway it cannot be recovered. But to assess whether information is lost or only concealed we need some theoretical results and to this we must now turn.

Viewed in this way, the retina is a device for sampling the visual image at intervals which, from the point of view of hyperacuity, might appear somewhat coarse. Is it possible to sample a continuous distribution and yet not lose any information? The mathematical answer to this is well-known. Provided that the pattern does not change too abruptly in space the reconstruction can, in theory, be perfect.

To make this more precise consider first the one dimensional case. The condition imposed is that, if the pattern is analysed into its Fourier components, they must be zero above a certain limiting spatial frequency. For any general spatial pattern we can achieve this by passing it through a perfect low-pass filter (see Figure 2). This allows all spatial frequencies below the cutoff frequency of the filter to pass unaltered while reducing to zero all frequencies above this limit. The sampling theorem then says that for such a pattern we need only sample it at the (regular) intervals shown in Figure 3. That is, the sampling points must be spaced no further apart than the zero's of the higher frequency spatial component, as shown in the figure. The proof is elementary (see the Appendix).

Moreover there is a relatively simple method of reconstructing the continuous distribution from the samples. This involves convolving the amplitudes of the sampled points with the mathematical function $\text{sinc } x = \sin \pi x / \pi x$. This function is shown in Figure 3.

The operation of convolution is described in the Appendix. In very rough terms, if one function is convolved with a second one, the resulting function would be considered as the first function spread everywhere by the second one (or vice versa). Thus, if the second function is a Gaussian, the effect of convolving it with a more extended function is to average the latter locally on every point and thus make it everywhere smoother.

We shall need one other useful result. This is Logan's zero-crossing theorem

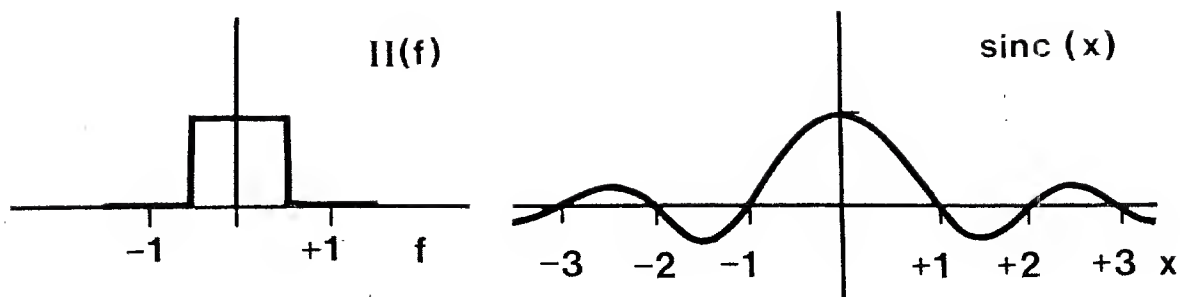


Figure 2. A perfect low pass filter is shown on the left. Frequencies above $|f| = 1/2$ are blocked, whereas frequencies below are passed undistorted. The inverse Fourier transform $\text{sinc } x$ is outlined on the right. The function sinc has an infinite number of side-lobes, not shown here, but indicated in Figure 3. Convolution of a one-dimensional pattern with the "receptive field" $\text{sinc } x$ is equivalent to filtering it with $\pi(f)$.

(Logan, 1977). Again we consider the one-dimensional case. This time we impose, not just an upper frequency limit to the filter but also a lower limit as well, to give a bandpass filter just one octave wide (see Figure 4). That is, we remove both the high frequencies and the low frequencies. This necessarily means that the filtered distribution must cross the zero line fairly often since there are no low-frequency components to keep it on one side of zero for any considerable distance. Logan's theorem states that it is almost always possible to reconstruct the entire (filtered) distribution, given only the positions and signs of the zero-crossings, subject to a constant multiplication factor. The exact conditions under which this can be done are set out in the Appendix. A similar theorem applies to the zero-crossings of a two-dimensional distribution (see the Appendix).

However in this case the proof is an existence proof. There appears to be no very simple way to reconstruct the distribution from the zero-crossings. Nevertheless the theorem shows quite clearly that, provided the bandpass filter is only one octave wide, the zero-crossings alone are a very rich source of information. It would therefore not be surprising if the brain used them as an important way to transmit and further process visual information, especially since the positions of the zero-crossings of a function are unaltered if the amplitude scale is altered, i.e. if the function is multiplied by a constant factor. This feature of the zero-crossings is an important aspect of Marr and Poggio's theory of stereopsis (1977, 1979).

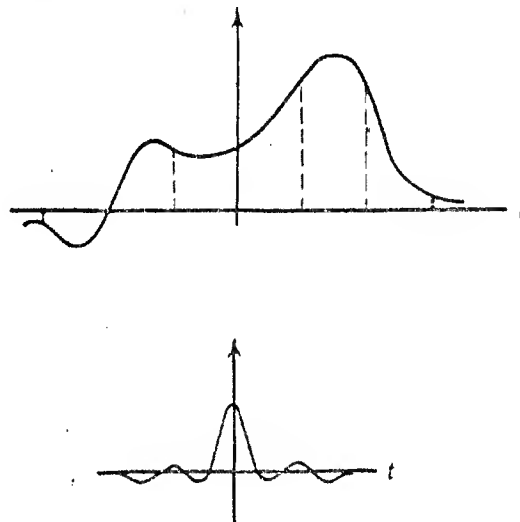


Figure 3. An illustration of the sampling theorem. The time function shown above is sampled at the Shannon rate. Convolution of the series of sampled values with the function $\text{sinc } x$ shown below reconstructs the original function exactly.

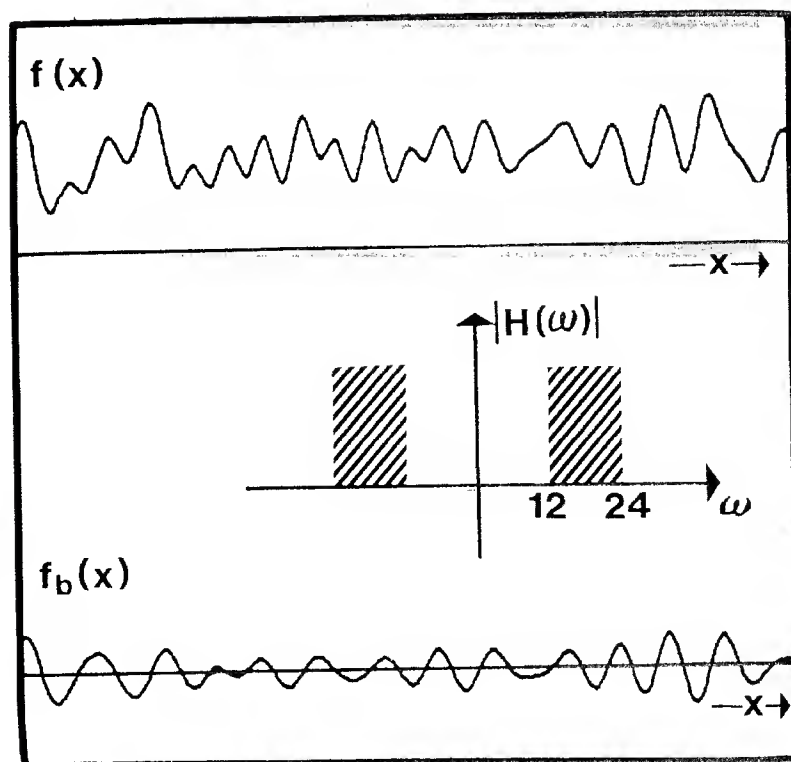


Figure 4. The meaning of Logan's theorem. A band-limited signal $f(x)$ is filtered through an ideal one-octave filter (inset) providing $f_b(x)$. Since $f_b(x)$ has a bandwidth of one octave and it has no complex zeros in common with its Hilbert transform and no multiple real zeros, Logan's theorem implies that f_b is determined, up to a multiplicative constant, by its zero-crossings alone.

4. The Optimum Filter

We must next consider what mathematical operations, in a very general sense, are likely to be performed on the intensity distribution of the retinal image as it proceeds along the visual pathway. It is well known that the ganglion cells of the retina hardly respond to an increase in general level of uniform intensity but rather to inequalities in it, in a "center-surround" manner. This is even more true for cells of the lateral geniculate and for cells in layer 4C of striate cortex of the monkey which are the main recipient of the input from the LGN (the cat's cortical cells may be somewhat different).

A useful way to approach this problem from a computational point of view is to ask what operation it would be best to perform on the visual input to the retina. Simply on computational grounds an important preliminary operation in the processing of visual information is the localization of sharp changes in image intensity, on the ground that these usually correspond to physically important items like edges in the image. A major difficulty with natural images is that changes can and do occur over a wide range of scales. It follows that one should seek a way of dealing separately with the changes occurring at different scales, since no single filter can be optimal at all scales. The appropriate filter, at each scale, should be approximately bandpass so that we reduce the range of scales over which intensity changes take place. The fact that there appear to be bandpass channels in vision would in any case point in this direction. In addition, since the "high" frequencies in the spatial pattern are filtered out we can use sampling techniques. As explained in Marr and Hildreth (1979) it is sensible to choose a function which is compact both in space (because the visual world is largely made up of compact features) as well as in frequency. The function which does this best is the Gaussian, since its Fourier transform is also a Gaussian.

The first (spatial) differential of an edge has a maximum, but the second differential has a zero-crossing at the point where the edge is located. In fact, an intensity change corresponds to a zero-crossing in the second spatial derivative. Thus we need to take the second differential of the image filtered through a Gaussian. This is equivalent to convolving the image with the second differential of a Gaussian. As shown in Figure 5 this is indeed a function which gives a center-surround type of visual field. A related function (indicated as $\nabla^2 G$) can conveniently be used for a two-dimensional spatial pattern. A similar function, the difference of two Gaussians, has been suggested by Wilson and Giese (1977). The details are given in the Appendix.

The Fourier transform of Figure 5 is shown in the same figure. As can be seen,

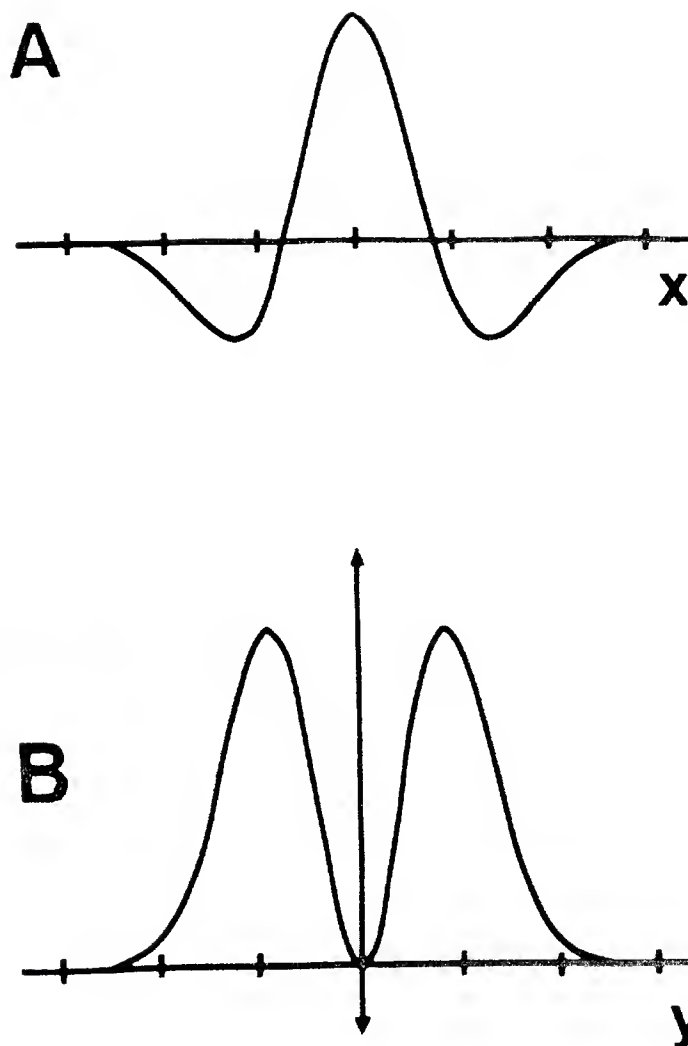


Figure 5. The 1-D $\nabla^2 G$. The corresponding two-dimensional field is circularly symmetric with a Mexican hat shape. B is the Fourier transform of A.

this is not exactly a bandpass filter of width one octave, but it approximates it in that both low and high frequencies are much diminished. The "effective band width" might be considered to be about an octave and a half. Thus it is unlikely that the whole (filtered) distribution could be recaptured from the zero-crossings alone (see Appendix) but nevertheless the positions of the signed zero-crossings probably contain a good fraction of the information in the continuous distribution. An example at one scale of the signed zero-crossings of a filtered two-dimensional image is given in Figure 6.

In summary our thesis is that the set of zero-crossings of the image filtered through independent $\nabla^2 G$ filters of 3 or 4 sizes (in order to cover the range of scales characteristic of natural images) represent the main "symbols" on which later visual processes, like stereopsis, are likely to operate.

Finally, observe that the physiological detection of zero-crossings need not depend on the detection of cells with zero-response. For instance, near an intensity edge the zero-crossings in the bandpass signal are flanked by two peaks of opposite sign. Detection of zero-crossings can thus be performed on the basis of peaks rather than zero-response. Marr and Hildreth (1979) and Marr and Ullman (1979) have proposed physiological schemes for how simple cells in area 17 may detect and represent oriented zero-crossing segments.

5. The Striate Cortex

The main input to the striate cortex is to layer 4C. To be more precise, in the monkey the major ganglion cell type of the retina, the X cells, which project to the parvocellular layers of the LGN, project from there mainly to 4C β of the striate cortex. The Y cells, though larger in size, are much fewer in number. They connect to the magnocellular layers of the LGN and from there project mainly to layer 4C α . Y cells respond more transiently than X cells. X cells give a fairly sustained response and are more linear in their behaviour. Here we shall mainly be considering X cells and cortical layer 4C β . For this reason our theory will be a linear one, even though linearity may be only a first approximation to the truth.

What immediately strikes one in examining layer 4C β is the very large number of very small cells in this layer. Thus, as Barlow (1979) has already suggested, it makes sense to consider that the visual image, or to be more precise a filtered version of it, is reconstructed there explicitly from the sampled input passed to it, via the LGN, from the ganglion cells of the retina. Barlow speaks loosely of "reconstructing" the visual image. Because of our computational theory of early visual information processing we

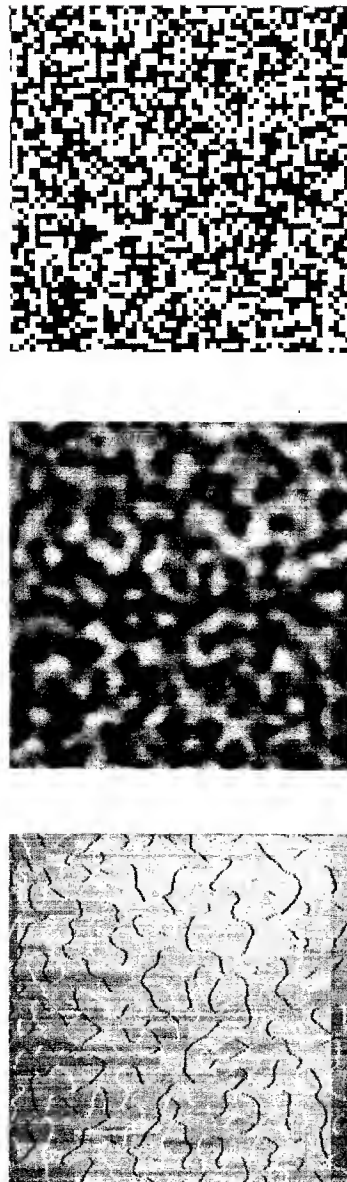


Figure 6. A pattern (a) is filtered (b) through a medium size $\nabla^2 G$ receptive field. Black areas represent negative values, white positive. Figure 6(c) shows the zero-crossings of (b).

would prefer to express this as a more precise hypothesis: that the cells of layer 4C β represent the reconstruction on a very fine grid of the visual image passed through a $\nabla^2 G$ filter in such a way that its zero-crossings are especially well preserved.

The different sizes of the $\nabla^2 G$ operator, required by our previous scheme, may correspond to various receptive field sizes of the X cells (at any given eccentricity) in the LGN. If this is so, it seems likely that the interpolation in layer 4C β mainly operates on the smallest of the X channels with a foveal central width of around 1.5' and a sampling density between 1' and 30" (Marr and Hildreth, 1979).

Exactly how the reconstruction might be done in layer 4C β is not clear at the present time. A major problem is how to represent the negative parts of a function. If all the cells of 4C β had a steady background firing rate, then a positive value could be represented by an increase in firing rate and a negative value (relative to the mean) by a decrease — or vice versa. Alternatively, if the resting firing rate were very low, then one distribution of cells could map the positive part of the function (being silent for the negative parts) and another, somewhat separate one, would map the negative parts. That is, in this latter set of cells a high rate of firing would represent a large negative value of the function.

Even if it were known which of these two alternatives were correct, there would still remain the problem of how to interpolate from the relatively sparse input. How would the reconstruction function be implemented? It may not be necessary to do more than provide the central positive peak of this function (ideally $J_1(\rho)/\rho$, see Appendix) and a surrounding rather shallow negative peak. The very small oscillation in amplitude beyond that could probably be ignored (see the Appendix). It is known that the axonal trees of the geniculate input spread out somewhat. Could the synapses in the inner part of the axonal tree excite, whereas those of the outer portions inhibit? This would do the job very nicely but we know of no precedent for such behavior.

A more plausible way to implement the reconstruction function would be by the sum of two Gaussians of opposite sign, a narrow positive one plus a lower, wider negative one, the total integral of one being equal and opposite to that of the other. Exactly how this would be done depends upon how the negative is represented, as discussed above.

More elaborate schemes are possible but are not without their difficulties. An alternative way, if there are indeed two "maps", one for the positive values of the function and the other for the negative ones, is to have a rather sharply localized inhibition of one map by the other. This would have the effect of sharpening up the

smearing produced by axonal spread and imprecision of wiring. Though we have yet to do computer simulation of this it seems probable that in most cases this would leave the position of the zero-crossing almost unaltered though the non-zero parts of the function might be distorted somewhat. It is even possible that, in this case, it might only be necessary to implement the central positive part of the reconstruction function. Preliminary computer studies show that the location of the zero-crossings can be reconstructed with Vernier precision for 1-D functions via a much simpler receptive field than the ideal $\text{sinc } x$ (Hildreth, 1980).

6. Experimental Evidence

Some of the experimental evidence has already been outlined by Barlow (1979). It should be remembered that most of the psychophysical data is obtained from human beings, whereas the greater part of the detailed neurophysiological and neuroanatomical data is from the macaque monkey. As it is believed that the visual system of man and the macaque, at least in its earlier stages, are not very different, no great harm is likely to be done by combining data from these two sources. We hope to give a detailed account of the numbers involved elsewhere. Here we merely sketch the broad features.

For the macaque the number of ganglion cells in one retina is about 1.5×10^6 and it is believed that the number of cells projecting to the striate cortex from the LGN is about the same (S. LeVay, personal communication and Le Gros Clark, 1941). Since the area of the striate cortex in one hemisphere is about 1400 mm^2 (Le Gros Clark 1941, 1942; Cowey 1964) there are, in all, about 500 incoming axons of each type from the LGN per square millimeter of cortex.

The total number of cells per square millimeter of striate cortex has been estimated at 3.5×10^5 . The exact fraction of these in layer $4C\beta$ does not appear to have been reported. A reasonable estimate is 10% giving 3.5×10^4 per mm^2 (Powell, personal communication). Thus there are about 50 times more cells in Layer $4C\beta$ than incoming LGN axons. Barlow (1979) arrived at a similar ratio (using Garey's data). These figures can only be regarded as very approximate but they make the general point. There are many cells in $4C\beta$ and this does indeed suggest that they might be used to make a fine-grained reconstruction of the sampled input provided by the incoming optic radiation.

If there are indeed about $3.5 \times 10^4 / \text{mm}^2$ then, lumping them all together, we see that their mean spacing is about 4 microns when projected onto the surface of the cortex. To translate this into a visual angle we need the so-called magnification factor.

This is of course different in different parts of the visual field. Near the fovea it is about 0.25° /mm. for the macaque (Daniel and Whitteridge, 1961). Ignoring possible complications produced by ocular dominance columns this implies that a spacing of 4μ corresponds to a visual spacing, near the fovea, of roughly $3''$ of arc. This is indeed fairly close to the observed hyperacuity limit for humans.

The above calculation can only be regarded as approximate, both because some of the numbers need to be determined more accurately and because the argument has been oversimplified. For example, we have ignored the fact that not all ganglion cells are X cells, that there might be two distinct maps (possibly more) and so forth. A similar calculation needs also to be done for the simple cells of the striate cortex, but even a rough estimate suggests that, for any one orientation, there are far fewer of them than the non-oriented cells of layer $4C\beta$.

If there is a rather precise reconstruction of the filtered visual input in layer $4C\beta$ then this should show up in single-electrode experiments. We are informed by Dr. David Hubel (personal communication) that preliminary results suggest that the mapping of the visual field in this layer is indeed rather precise. It would be useful to have an estimate of just how accurate it really is. The anatomical mapping between LGN and layer $4C\beta$ has in fact to be quite precise. Preliminary calculations show that the ordering of the LGN inputs has to be preserved, i.e. the jitter in the LGN inputs must be less than $30''$ (in the fovea) in order to ensure a reconstruction at Vernier accuracy. Our hypothesis would furthermore require that the cells in layer $4C\beta$ should have the same receptive field as the corresponding LGN cells and this seems again to be supported by known physiological data. In addition any detailed theory must take into account neuroanatomical factors such as the spread of the incoming axonal trees. For axons from the parvocellular cells of the LGN this is believed to be around 500μ (S. LeVay, personal communication). However, it should be remembered that this is the total spread. The parameter used in the mathematics, which assumes that the axon terminals have an approximately Gaussian distribution, is σ , the "half-width" of the Gaussian (see the Appendix). This may well be less than 100μ , corresponding, in the foveal representation, to about $1'$. Whether this degree of smearing at the input can be satisfactorily sharpened, and by what means, remains to be seen.

As more becomes known about the neuroanatomical details of the visual cortex it may be possible to suggest more precisely how an exact reconstruction could be implemented. More quantitative data would be especially welcomed, especially that which could be obtained fairly exactly. We know no theory yet which justifies our asking for cell counts, etc., to be as accurate as 10% or even 20%, but in some cases the relevant numbers are not known to a factor of 10 (for example, how many cells in

the striate cortex project back to the LGN?). A rough estimate, to within a factor of two is always better than no estimate at all. A factor of about $\sqrt{2}$ might be the sort of precision to aim for at this stage.

7. Conclusion

Our hypothesis is that hyperacuity depends on a fairly accurate reconstruction of the filtered visual input, with well-preserved zero-crossings, in layer $4C\beta$ of the striate cortex. As Barlow has already argued (1979) this seems plausible enough as a working hypothesis, but, as he has pointed out with especial force, it may seem to suffer from one major disadvantage: it is too static. Recent psychophysical results, especially those of Burr (Burr and Ross, 1979 and Burr, 1979) have emphasized what was known before: that it is not possible to understand hyperacuity without considering the response to moving objects. Whether this involves the Y cells of the retina, which project to the cortex mainly to layer $4C\alpha$, or whether the effect of movement depends also on the W ganglion cells of the retina, projecting to the superior colliculus and from there to the striate cortex via the pulvinar, or whether other cortical areas, such as Zeki's movement area on the posterior bank of the superior temporal sulcus (Zeki, 1974) are especially involved — all these questions remain for the future. It is pointed out in the Appendix that, from the point of view of communication theory, psychophysical experiments of the Burr type may not present a problem for our proposal. They may be satisfactorily explained by a scheme in which the zero-crossings of the filtered image — but not necessarily the filtered image itself — are precisely reconstructed on the fine grid of layer $4C\beta$. This issue can be resolved only by additional psychophysical experiments. How such computations might be implemented in the brain is not immediately obvious.

Though some clues may come from neuroanatomy we cannot help feeling that little progress will be made until the response to stationary and moving spots of light, of the type now being studied on man by Burr, is also measured in the alert monkey with electrodes in various regions of the brain; for instance, do cells in layer $4C\beta$ reconstruct the pattern of activity at moments intermediate between the flashes and at locations intermediate between the stations at which the line segments are flashed?

APPENDIX

1. Mathematical Tools and Definitions

The convolution of two functions $f(x,y)$ and $g(x,y)$ is defined as the function $h(x,y)$

$$(1) \quad h(x,y) = f(x,y) * g(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u,v) g(x-u, y-v) du dv$$

Figure A1 illustrates graphically the meaning of this operation for one-dimensional functions. The importance of convolution stems from the fact that the output of a linear time-invariant system on an input function is given by the convolution of the input with the impulse response characteristic of the system. Thus, the operation of linear filtering (in space as well as in time) is equivalent to convolution.

The Fourier transform $\tilde{g}(f_x, f_y)$ of the function $g(x,y)$ is defined here as

$$(2) \quad \tilde{g}(f_x, f_y) = F\{g(x,y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) \exp(-j 2\pi (f_x x + f_y y)) dx dy$$

The transform is itself a complex-valued function of two independent variables f_x and f_y , which we generally refer to as frequencies.

2. The Sampling Theorem

The retinal image is sampled at a set of discrete points by the photoreceptors and it is represented by a still smaller number of nerve fibers in the optic radiation. Intuitively, it is clear that if these samples are taken sufficiently close to each other, the sampled data provide an accurate representation of the original 2-D function, in the sense that the pattern can be reconstructed with considerable accuracy by simple interpolation. For band-limited functions the reconstruction can be accomplished exactly, provided only that the interval between samples is not greater than a certain limit. The retinal image is effectively band-limited by diffraction at the pupil to about 60 cycles per degree in man.

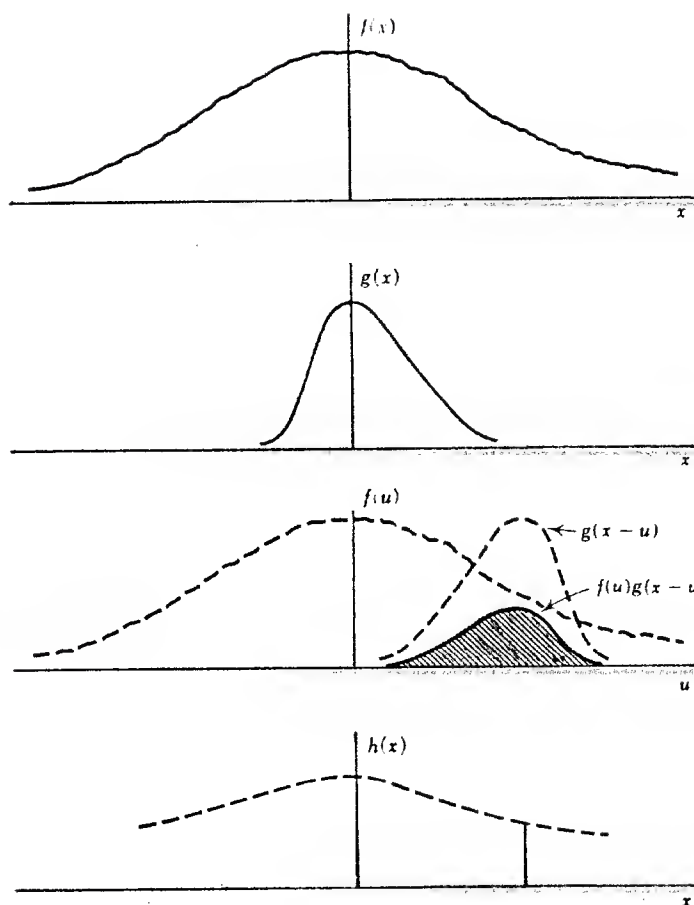


Figure A1. A pictorial illustration of the convolution operation. The function $f(x)$ is convolved with $g(x)$ according to $h(x) = \int_{-\infty}^{\infty} f(u) g(x-u) du$. The value $h(x)$, represented by the height of the segment in the bottom diagram equals the shaded area above.

We derive here (following Goodman, 1968) a simple version of the sampling theorem for 2-D functions (possibly time-dependent). We will also show how the conditions of the classical sampling theorem can be relaxed, if the pattern to be sampled is known to move at constant (known) speed. Sampling in time will not be considered; it is easy to extend the proofs to this case.

3. The Sampling Theorem in 2-D

Let us consider a rectangular lattice of sampling points as shown in Figure A2. With the continuous function $g(x,y)$ we associate its sampled version

$$(3) \quad g_s(x,y) = g(x,y) \text{comb}(x/X) \text{comb}(y/Y)$$

where $\text{comb}(x) = \sum_n \delta(x-n)$. The sampling function $\text{comb}(x/X) \text{comb}(y/Y)$ consists of an array of δ functions, spaced at intervals of width X in the x direction and width Y in the y direction. The Fourier transform of $\text{comb}(x/X)$ is

$$(4) \quad F\{\text{comb}(x/X)\} = \sum_n \delta(f - n/X)$$

as shown in Figure A3.

The Fourier transform of g_s is

$$(5) \quad \tilde{g}_s(f_x, f_y) = XY \text{comb}(f_x X) \text{comb}(f_y Y) * \tilde{g}(f_x, f_y)$$

Thus, the spectrum of g_s can be found simply by replicating the spectrum of g about each point in the f_x, f_y plane corresponding to the lattice defined by $\text{comb}(f_x X) \text{comb}(f_y Y)$, as shown in Figure A4.

The function g is assumed to be band-limited and thus its spectrum \tilde{g} is nonzero over only a finite region of the frequency space, called its support. If the sampling points are sufficiently close together (i.e. X and Y are sufficiently small), then the separations $1/X$ and $1/Y$ of the various "lobes" in Figure A4 will be great enough to

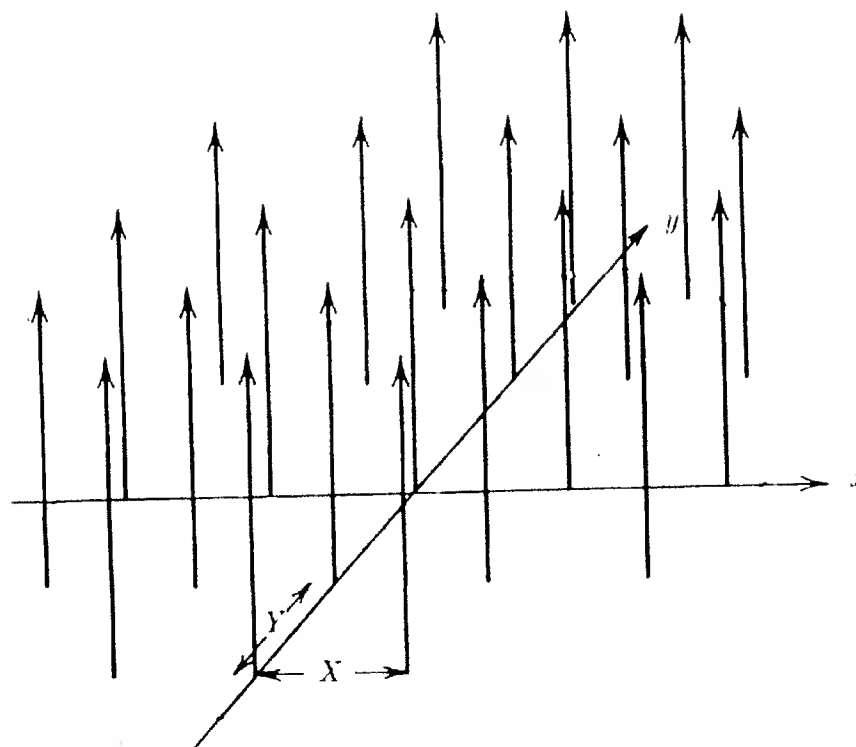


Figure A2. The "sampling function" $\text{comb}(x/X) \text{comb}(y/Y)$. By multiplying a continuous function $g(x,y)$ with this array of delta functions one obtains the sampled function g which essentially consists of the values of the original function at the positions of the arrows.

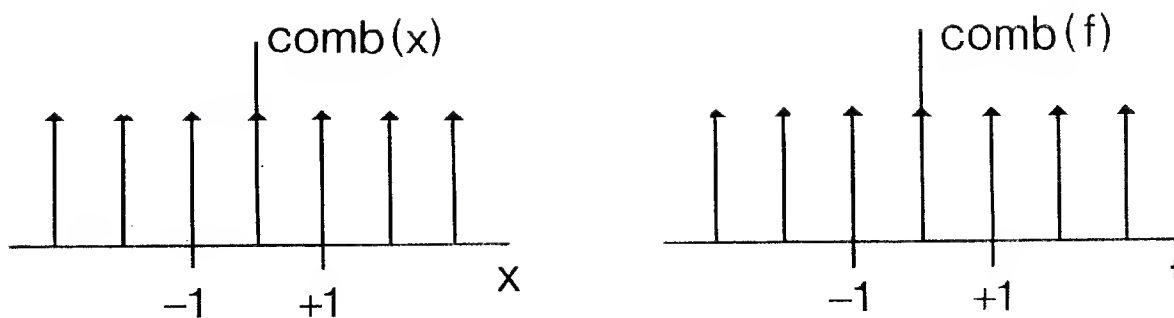


Figure A3. The function $\text{comb } x = \sum_n \delta(x - n)$ and its Fourier transform $\text{comb}(f) = \sum_n \delta(f - n)$. The comb function is thus its own Fourier transform. The small ticks show where the variable has a value of unity.

ensure that adjacent regions do not overlap. We can reconstruct $g(x,y)$ from \tilde{g}_s if we can recover \tilde{g} from \tilde{g}_s , and this can be accomplished exactly by passing \tilde{g}_s through a linear filter that excludes all side-lobes in \tilde{g}_s , but includes the central one. In the limiting case, adjacent regions in the spectrum \tilde{g}_s just touch. This happens (see Figure A4) when $X = 1/2B_x$ and $Y = 1/2B_y$, where $2B_x$ and $2B_y$ represent the width in the f_x and f_y directions, respectively, of the smallest rectangle that completely enclose the support of \tilde{g} (support centered on the origin). In this case the form of the filter could be

$$(6) \quad H(f_x f_y) = \pi(f_x/2B_x) \pi(f_y/2B_y)$$

$$\text{where } \pi(x) = -1 \quad |x| \leq 1/2$$

$$0 \quad \text{otherwise.}$$

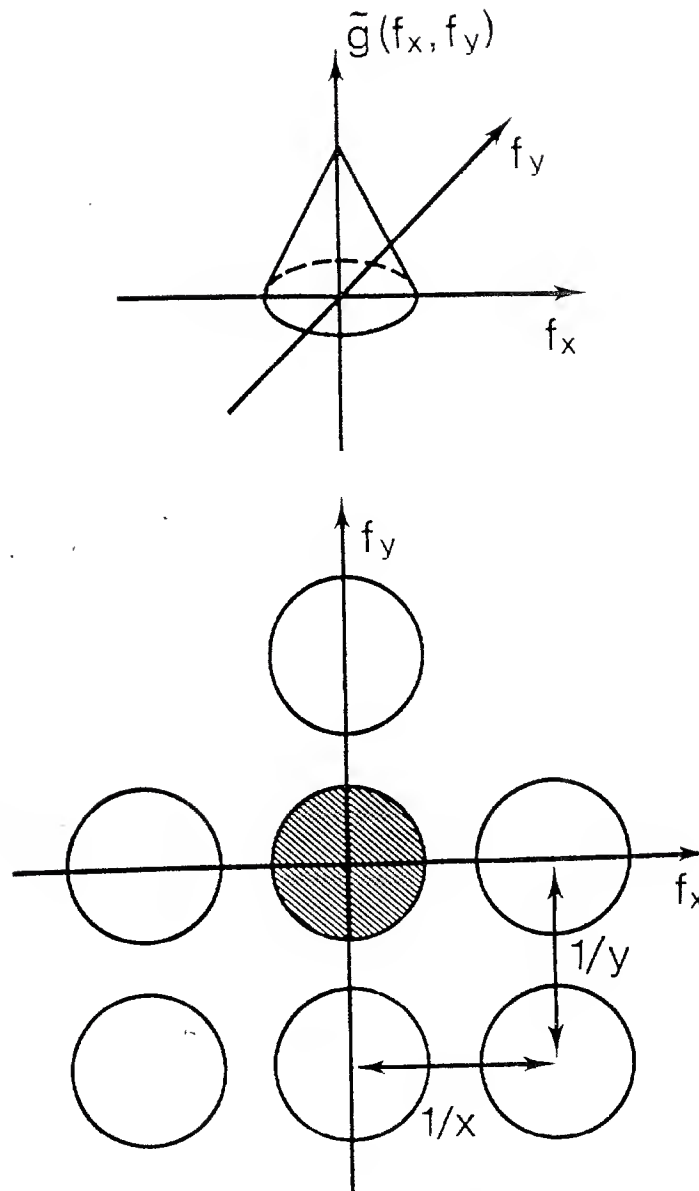


Figure A4. The Fourier transform \tilde{g} of a band-limited pattern $g(x,y)$, chosen as an example to be a cone, and (b) the support of the Fourier transform \tilde{g}_s of the sampled function

$$g_s(x,y) = g(x,y) [\text{comb}(x/X) \text{comb}(y/Y)]$$

where the expression within brackets is the sampling function shown in Figure A2. The spectrum \tilde{g} is repeated at distances $1/X$ and $1/Y$. If the sampling distances X and Y are small enough compared with the bandwidth of g , there is no overlap of the side-lobes in the spectrum of g_s . Only some of the side-lobes are shown here: they repeat over the whole Fourier plane.

Filtering \tilde{g}_s through H one gets

$$(7) \quad \tilde{g} = \tilde{g}_s(f_x f_y) H(f_x f_y)$$

which corresponds, as it can be seen by Fourier inversion, to the following (classical) interpolation scheme

$$(8) \quad g(x, y) = XY \sum_{n, m} g_s(nX, mY) \text{sinc}[2B_x(x - nX)] \text{sinc}[2B_y(y - nY)]$$

where $\text{sinc}(x) = \sin \pi x / \pi x$.

There are, of course, especially in 2-D, a variety of filters that could perform a correct interpolation. For instance, another choice for H is (if $B_x = B_y$)

$$(9) \quad H(f_x f_y) = \text{circ}(B_x)$$

$$(10) \quad \text{where } \text{circ}(r) = \begin{cases} -1 & \text{if } r \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

as shown in Figure A5, with an inverse Fourier transform

$$(11) \quad F^{-1}\{\text{circ}(r)\} = J_1(2\pi\rho)/\rho$$

J being the Bessel function of the first kind, order one. Figure A5 illustrates the circle function and its transform. Filtering \tilde{g}_s through $\text{circ}(B_x)$ corresponds to convolving g with a "receptive field" of the type shown in Figure A5(b).

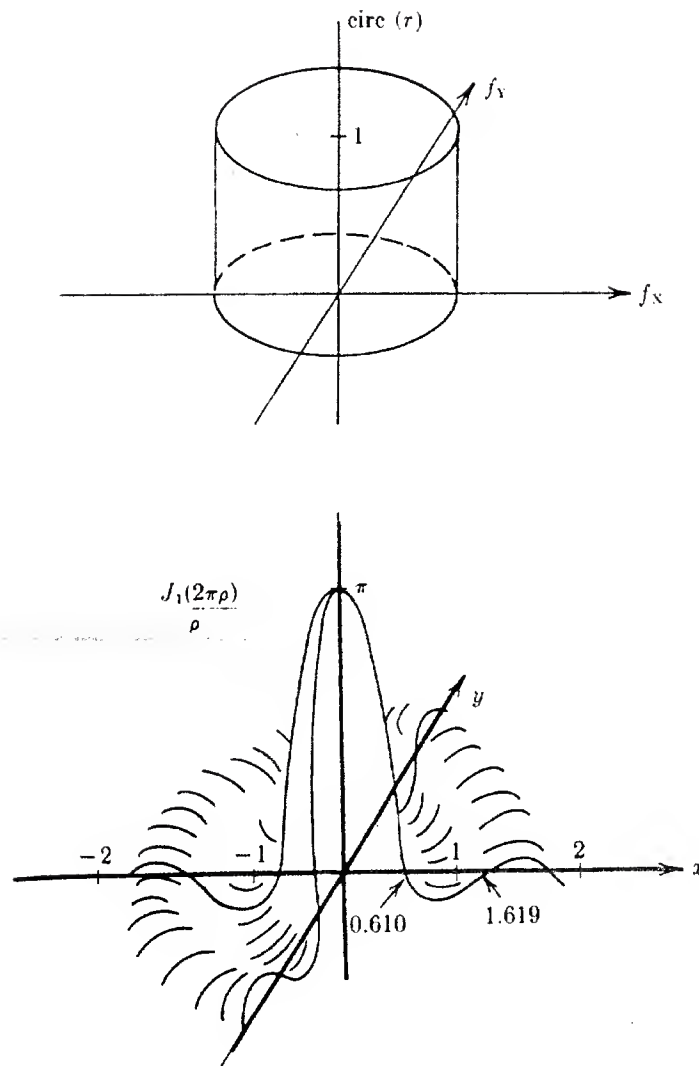


Figure A5. (a) The ideal low pass filter circ (r) in the Fourier plane and (b) its inverse Fourier transform, i.e. the corresponding receptive field.

We have shown the following:

Theorem 1 (Whittaker, Shannon): A band-limited function $g(x,y)$ can be recovered exactly from a rectangular array of its sampled values, according to eq. 8. The sampling distance in the x or y direction can be as large as the reciprocal of twice the bandwidth on f_x or f_y .

This proof of the sampling theorem can be easily extended to a non-rectangular sampling lattice, for instance a more efficient (and more realistic, for the human retina) hexagonal one. Since the retinal image is band-limited by the optics to about 60 cycles per degree, Theorem 1 for a hexagonal lattice implies that the maximum distance between photoreceptors should be $\Delta\phi = 1/60 \sqrt{3} = 27$ sec., which is about right for the human fovea. The finite diameter of the photoreceptors worsens only slightly the overall transfer function of the system. The proof clearly shows that the classical interpolation function sinc does not have an exclusive role. Many other interpolation functions would do as well or almost as well, especially if the sampling density is higher than the minimum. In such a case, the sidelobes in the spectrum \tilde{g} , do not touch; thus the requirement of the sharp cutoff in the filter H is removed. A variety of filter functions would exclude the sidelobes while transmitting the central lobe without significant distortions.

With a suitably higher sampling density, it is clear that simple receptive fields (see for instance Figure 5) could interpolate almost as well as $J_1(2\pi\rho)/\rho$ especially with respect to the location of the zero-crossings (Figure 6). Recent computer experiments show indeed that very simple interpolation schemes (linear interpolation, filtering through a Gaussian, filtering through $\nabla^2 G$) can localize zero-crossings in the output of a channel having $w = 1' 30''$ with vernier precision (E. Hildreth, 1980).

4. The Sampling Theorem for Bandpass Functions

Stronger results than Theorem 1 hold if the band-limited function is also bandpass. For instance, a one-dimensional bandpass function with one octave bandwidth can be sampled at half the rate set by the classical sampling theorem. This can be easily seen in Figure A6(a). In 2-D, sampling below the classic limit corresponds to interlacing of the lobes in the Fourier spectrum of the sampled function. Depending on the geometry of the bandpass support, interlacing can take place without overlapping of the side-lobes. Suitable filtering can again recover exactly the original function (Figure A6(b)). Notice that this scheme cannot be applied to a 2-D bandpass

function with a circularly symmetric, ring-like support in the Fourier plane.

However, the (sufficient) conditions given by this overlap argument are not strictly necessary in the bandpass case. It turns out that, despite eventual overlap, it is still possible to reconstruct exactly the bandpass function with two interlaced sampling sequences, each one having a sampling interval of $1/B$, where B is the width of the band (for $f > 0$). The average sampling interval is thus $1/2B$.

The following theorem holds:

Theorem 2: A bandpass function can be recovered exactly from a rectangular array of its sampled values, through suitable filtering. The sampling distance on the x or y direction can be as large as $1/2B$, where B is the width of the non-zero part of the spectrum (for positive frequencies only).

5. The Sampling Theorem For Moving Patterns

For simplicity we consider a 1-D pattern $f(x)$, band-limited in spatial frequency, sampled at a regular 1-D array of points. Movement of the pattern $f(x)$ produces a function $f(x, t)$ whose sampling in space obeys the usual restrictions set by the sampling theorem. If we assume, however, that $f(x)$ moves at a constant (known) velocity v , it is rather clear that the sampling rate may become very low, without losing information (in fact one "photoreceptor" clearly suffices). We analyze this problem with the same methods as in the previous part. The proof is quite instructive, especially for situations in which the velocity is not exactly known.

Since $g(x)$ moves at constant speed v the pattern of excitation on the "retina" is

$$g(x, t) = g(x - vt).$$

The Fourier transform is

$$(12) \quad \tilde{g}(f_x, f_t) = \tilde{g}(f_x) \delta(f_t - v f_x)$$

where $\tilde{g}(f_x)$ is the Fourier transform of $g(x)$. Since $g(x)$ is band-limited, $g(x, t)$ is

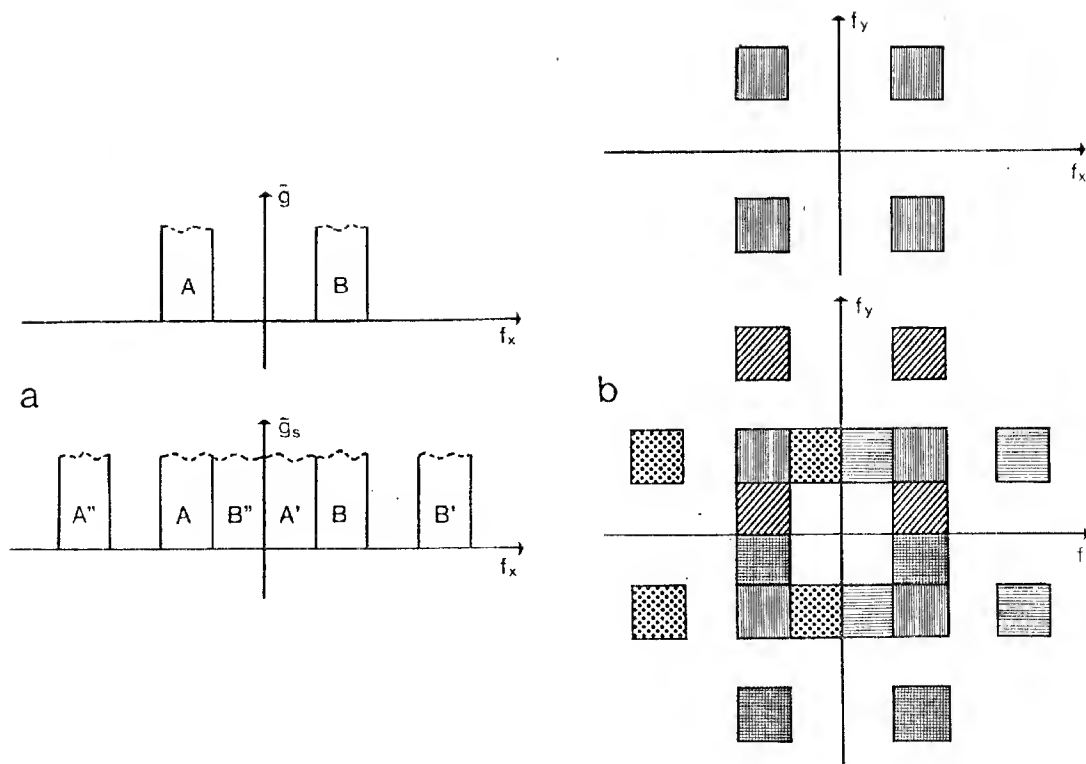


Figure A6. (a) \tilde{g} is the (modulus of the) Fourier transform of a one-octave bandpass signal. If the signal is sampled at half the Shannon rate, interleaving of side-lobes but no overlapping occurs in the spectrum \tilde{g}_s . \tilde{g}_s is obtained by shifting g by all integral multiples of Bf_x , Bf_x being the bandwidth of g . Only two of the side-lobes are shown in the bottom of (a). The same hatching identifies parts of the same side-lobe.

(b) The same situation for a two-dimensional pattern assumed to be bandpass (one-octave) on f_x and f_y . Again sampling at half the classical rate corresponds to interleaving but no overlapping of the side-lobes. Only some of the side-lobes are shown in the lower half of (b). They are created by shifting the original spectrum (shown in the upper half) by all multiples of $1/X$ and $1/Y$, X and Y being the sampling distances. Side-lobes thus fill the whole Fourier plane.

Notice that if the support of g were a ring in the Fourier plane, overlap of the side-lobes in \tilde{g}_s would occur for any sampling rates lower than Shannon's rate.

band-limited in both space and time if v is finite. The array of "photoreceptors" spaced by X provides a sampled version $g_s(x - vt)$ whose spectral support is shown in Figure A7. The slope of the line is v . The classical sampling theorem requires a sampling distance $x \leq 2B_x$. In this case a filter function like $\pi(f_x/2B_x)$ can separate the central region from the side lobes and thus retrieve $g(x - vt)$. Figure A8 shows clearly, however, that the sampling distance X can be increased much above the classical limit irrespective of the velocity v , provided that v is different from zero: in the Fourier plane the distance between the side lobes can be made arbitrarily small (corresponding to X arbitrarily large in x space) without overlapping.

The original spectrum can be retrieved by the filter depicted in the Figure A8(a). The retrieval scheme in the limit of very large X requires convolution of $g_s(x, t)$ with the (noncausal) receptive field $\delta(x + vt)$.

Uncertainty in the velocity forces a finite sampling distance. The minimum allowed sampling rate depends on the geometry of the support of the Fourier transform. It can be derived easily from graphs like Figure A8.

For instance, assume that only the direction of motion is known, i.e. the sign of v . Figure A8(b) Shows that if only the sign of v is known the minimum distance on the $f_x f_t$ plane to ensure no overlapping is half the classical distance. Thus, the maximum distance between the samples can be twice the limit set by the classical sampling theorem. Again from the figure it is easy to derive the required filter and the corresponding "receptive field".

We have proved the following:

Theorem 3: Assume that a band-limited spatial pattern $g(x)$ with a bandwidth $2B_x$ moves at constant velocity v . Then, if the velocity v is known with arbitrarily high precision the distance between sampling points can be made arbitrarily large. Uncertainty in v requires a well defined maximum sampling distance, higher, however, than the classical limit (if $v \neq 0$). As a corollary, if only the sign of the velocity is known, the maximum sampling distance X is twice the classical limit for stationary patterns (thus $X = 1/B_x$).

This result can help in discussing the demonstration published very recently by Burr (1979). In his experiment, vernier line segments are displayed stroboscopically at a sequence of stations. Spatial offsets were detected with an accuracy in the 5 second range in spite of the movement (compare Westheimer and McKee, 1975); in addition an illusory displacement occurs if the line segments are accurately aligned in space, but

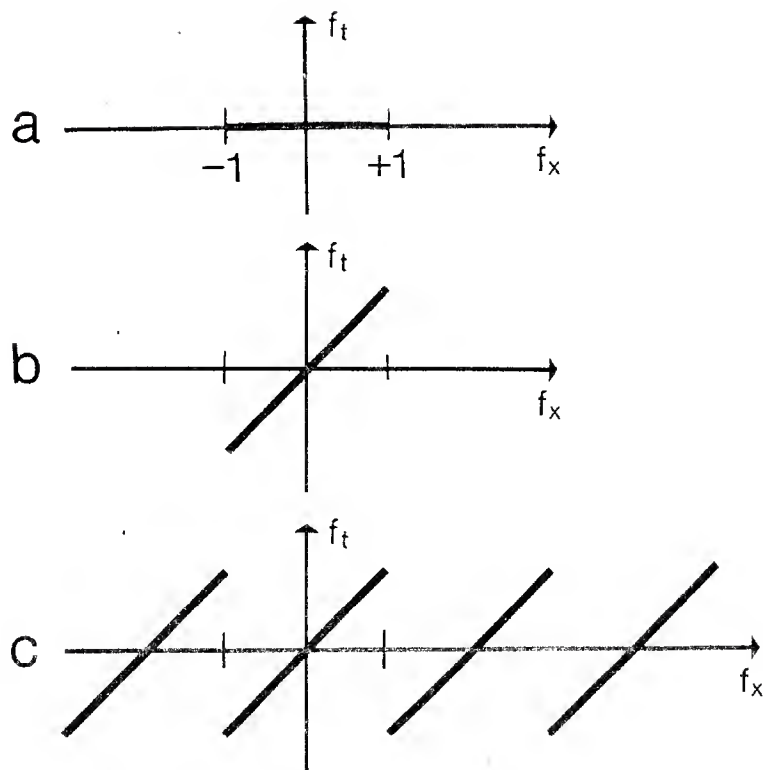


Figure A7. (a) The support (a "delta" segment) of the spectrum \tilde{g} of a band-limited one-dimensional pattern $g(x)$ at rest ($v=0$) in the Fourier plane of spatial (f_x) and temporal (f_t) frequencies. $g(f_x, f_t)$ equals $g(f_x)$ on that segment.

(b) The spectral support of the same pattern moving at constant speed v . The Fourier transform of $g(x-vt)$ takes nonzero values only on the "delta" segment, whose slope is v .

(c) Spectral support of $g_s(x, t)$. The pattern $g(x-vt)$ is sampled in space at the Shannon rate.

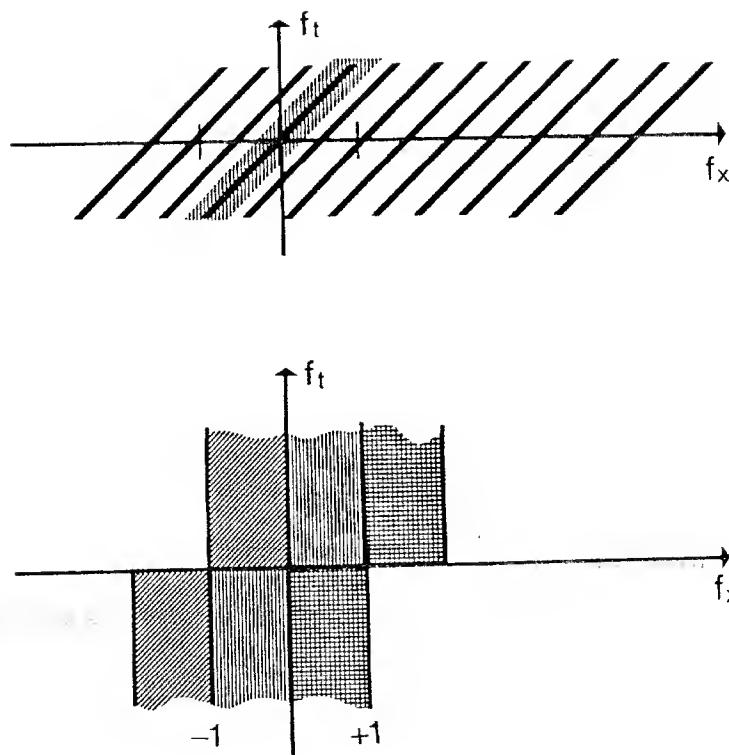


Figure A8. (a) The support of $g_s(x - vt)$, when the sampling rate is lower than Shannon's rate. The original function \tilde{g} (Figure A7(b)) can be retrieved by filtering the sampled data, for instance, with the filter indicated by hatching. The filter is arbitrary provided that it transmits without distortion the central lobe, eliminating at the same time all side-lobes.

(b) If it is only known that the sign of the velocity is positive, the support of \tilde{g} may lie anywhere in the region shown by vertical hatching. Sampling spatially at half the Shannon rate still ensures that the side-lobes do not overlap in the spectrum \tilde{g}_s . Compare Figure A7(c).

are displayed with a slight delay in one sequence relative to the other. The accuracy of detecting the equivalent displacement is again in the hyperacuity range. As Barlow (1979) pointed out, this suggests that the spatial pattern of activity at moments intermediate between the stroboscopic flashes is actually reconstructed. Clearly temporal interpolation (i.e. temporal low-pass filtering) followed by the "static" spatial interpolation can reconstruct a pattern of activity $g(x, t)$ on an arbitrarily fine spatial grid, from the sampled functions provided by the LGN fibers. This amounts to saying, as pointed out by S. Ullman, that almost nothing needs to be done in order to obtain the right kind of temporal filling-in. Temporal low-pass properties of the LGN pathway provide temporal blurring; spatial interpolation, for instance in layer 4C β , would then reconstruct activity between the LGN fibers at all times. Notice that the blurring due to temporal integration can be corrected at a later stage, for instance, by spatial high pass operations (for a pattern moving at constant speed, temporal and spatial variables are interchangeable). Thus, there is no problem in reconstructing the correct pattern of activity in 4C β for a real movement of the retinal image. Difficulties may appear, however, when motion of the object is simulated by presenting the image at discrete positions at separate instants, as in the experiments by Westheimer and by Burr. If the positions at which the vernier segments are flashed correspond to the sampling grid of the LGN cells, the simulated motion is in fact completely equivalent (from the LGN point of view) to real motion. If, however, neighbouring positions are much farther apart than neighbouring LGN sampling points and larger than LGN receptive fields, there may be too few samples (in terms of the classical sampling theorem) to reconstruct the equivalent spatio-temporal pattern of activity between positions. It is important to stress, on the other hand, that our hypothesis requires a precise reconstruction in space and time of the zero-crossings of the filtered image but not necessarily of the filtered image itself. Thus the necessary conditions — and we are indebted to S. Ullman for this remark — are probably weaker: computer experiments (by S. Ullman) show that the correct motion of a zero-crossing between stations can be obtained under conditions that forbid a faithful reconstruction of the corresponding 2-D function. From this point of view the bounds on the sampling intervals given by the theorems of this appendix are too strong, since they refer to a correct reconstruction of a whole function and not only of its zero-crossings.

Although more psychophysical experiments are absolutely essential to determine how the Vernier acuity measured by Burr actually degrades with increased separations between the positions at which the segments are flashed, it is indeed possible that separations larger than the maximum allowed by the sampling theorem may not dramatically reduce performance. Clearly these estimates depend also on which channel is actually reconstructed in layer 4C β : if larger receptive fields are involved, the maximum separation allowed would be correspondingly larger. For instance, for the

smallest channel ($w = 1' 30''$) the maximum distance should not be much larger than about $1'$. This estimate could become up to 8 times larger for the biggest (and probably transient) channel. It is also conceivable that the effective receptive field size of the channel reconstructed in layer $4C\beta$ may be transiently larger at the onset of a stimulus, either because of Y influences or via a mechanism similar to the one postulated by Detwiler et al. (1978) for the rod network in the retina. This would provide a lower cut-off in spatial frequencies for moving stimuli, allowing a larger sampling interval (see Theorem 1).

In any case, Theorem 3 shows how information about the velocity of an object moving at constant speed, could be used to increase the maximum (spatial) sampling interval (in principle Theorem 2 could also be exploited, because of the bandpass properties of the LGN channel). Thus, even rough and implicit estimates of the velocity may account, at least from the point of view of information theory, for eventual very large separations in Burr's type of experiments. For instance, it is conceivable that the interpolation scheme may be based on the a priori assumption of a movement of the retinal image within a well defined range of velocities. The correct motion may usually be provided by eye movements.

The sampling theorems outlined here do not, of course, say which neural mechanisms may be involved. But even if an explicit reconstruction on a finer grid of neurons is not done in our brain, these results characterize the conditions under which information of the Vernier type can be preserved in the visual pathway.

6. Logan's Theorem

If a one-dimensional band-limited function (belonging to B_∞ , i.e. the restriction to the real line of entire functions of exponential type, see Logan 1977)

a) has no free zeros (i.e. complex roots in common with its Hilbert transform) and no multiple real zeros,

b) is bandpass with enough real zeros in proportion to its bandwidth,

then the function is uniquely determined, up to an overall multiplicative constant by its (real) zero-crossings.

Condition (a) is almost always satisfied. Condition (b) is critical: it is always ensured if the bandwidth of the signal is less than one octave. For particular classes of

bandpass signals (b) may be satisfied even for larger bandwidths: for instance ergodic gaussian bandpass signals satisfy condition (b) if their bandwidth is less than 1.67 octaves (H.K. Nishihara, personal communication). On the other hand Logan's result is valid for ideal bandpass functions and it cannot be extrapolated with abandon to "almost bandpass" functions.

7. An Extension of Logan's Theorem to 2-D Functions

It is impossible to use directly Logan's technique for proving some 2-D version of its theorem. There is, however, a simple way of translating the 2-D problem into a 1-D problem in order to use Logan's result. In this way it can be shown that zero-crossings in a suitably bandpass 2-D function (in principle) determine the function within a multiplicative constant. The conditions under which this result is valid are likely to be too restrictive: they are sufficient but almost certainly not necessary. The argument runs as follows: the image $f(x, y)$ is filtered through one-octave bandwidth vertical masks: zero-crossings are then measured along horizontal scan lines at intervals appropriate to recover all the information. Since the one-dimensional functions associated with each scan line are then bandpass with less than one octave bandwidth, they satisfy Logan's theorem. The same operation can be carried out with horizontal masks and vertical scans. It can be shown that it is thus possible to reconstruct the filtered image (through horizontal and vertical marks) modulus a single scaling factor (Marr, Poggio and Ullman, 1979; Nishihara, in prep.; Poggio, in press). This reconstruction scheme cannot be applied to images filtered through non-oriented bandpass filters, like the circularly symmetric receptive fields of the ganglion cells.

8. $\nabla^2 G$

This way of locating zero-crossings in a filtered image is not the only nor necessarily the best method. It has been shown (Marr and Hildreth, 1979) that under certain rather weak conditions the zeros in an image filtered through a concentric type receptive field provide an equivalent way of locating edges, whose orientation must then be represented. This suggests that zero-crossings in an image filtered through concentric receptive fields may also contain all of the information of the image. As we mentioned, however, the extension of Logan's theorem to this case is not yet available.

In a similarly negative vein, we do not have yet any formal result on the information content of the zero-crossings in a non-ideal bandpass signal, although recent computer experiments (K. Nishihara, 1979) indicate that they still contain almost full

information, under rather loose conditions.

In summary, Logan's theorem cannot strictly be used in a theory of early visual processing; the important point is that it shows that zero-crossings of a bandpass signal are very rich in information. In this sense, it supports a set of computational arguments (Marr, 1976; Marr and Poggio, 1977, 1979; Marr and Hildreth, 1979) suggesting that the detection of zero-crossings in the output of independent (spatial) roughly bandpass channels is one of the first steps in the processing of visual information. Marr and Hildreth (1979) argue that intensity changes in the image at one scale may be detected by filtering (i.e. convolving) the image with the Laplacian of a two-dimensional Gaussian at that particular scale and then locating zero-crossings. Since the Gaussian is given (in 1-D) as

$$G(x) = 1/\sigma\sqrt{2\pi} \exp(-x^2/2\sigma^2)$$

its Laplacian (in one-dimension)

$$G''(x) = -1/\sigma^3\sqrt{2\pi} (1 - x^2/2\sigma^2) \exp(-x^2/2\sigma^2)$$

looks like a Mexican hat operator (see Figure 5), it is an approximately bandpass operator with a halfpower bandwidth of about 1.25 octaves and is very closely approximated by Wilson and Giese's (1977) difference of two gaussians (with a ratio for their σ 's of about 1.6). By using a range of $\nabla^2 G$ mask sizes one can deal with the wide range of scales over which intensity changes take place in a natural image.

These ideas may begin to account, on purely information processing grounds, for the presence of spatial-frequency-tuned channels in early human vision (Campbell and Robson, 1968; Wilson and Giese, 1977) and for the properties of simple cells in the cortex, which are usually described as detectors of edges and bars of various widths and orientations (Hubel and Wiesel, 1962).

ACKNOWLEDGEMENTS

We are grateful to H. B. Barlow and S. Ullman for discussions, to I. Geiss and A. Hartheimer for typing the manuscript and to L. Heimbürger for drawing the figures. F. H. C. Crick was supported by the Eugene and Estelle Forkhanf Foundation, J. W. Kieckhofer foundation and Samuel Roberts Noble Foundation, Inc. D. Marr and T. Poggio thank the Salk Institute for their kind hospitality.

REFERENCES

- Barlow, H. B. (1979) "Reconstructing the Visual Image in Space and Time," *Nature*, Vol. 279, pp. 189—190.
- Beck, J. and Schwartz, T. (1978) "Vernier Acuity with Dot Test Objects," *Vision Research*, Vol. 19, pp. 313—319.
- Burr, D. C. and Ross, J. (1979) "How does Binocular Delay Give Information About Depth?," *Vision Research*, Vol. 19, pp. 523—532.
- Burr, D. C. (1979) "Research Note. Acuity for Apparent Vernier Offset," *Vision Research*, pp. 835—837.
- Campbell, F. W. and Robson, J. (1968) "Application of Fourier Analysis to the Visibility of Gratings," *J. Physiol.*, Vol. 197, pp. 551—566.
- Cowey, A. (1964) "Projection of the Retina on to Striate and Prestriate Cortex in the Squirrel Monkey, *Saimiri Sciureus*," *J. Neurophys.*, Vol. 27, pp. 366—393.
- Daniel, P. M. and Whitteridge D. (1961) "The Representation of the Visual Field on the Cerebral Cortex in Monkeys," *J. Physiol.*, Vol. 159, pp. 203—221.
- Detwiler, P. B., Hodgkin, A. L. and McNaughton, P. A. (1978) "A Surprising Property of Electrical Spread in the Network of Rods in the Turtle's Retina," *Nature*, Vol. 274, pp. 562—565.
- Goodman, J. W. (1968) *Introduction to Fourier Optics*, McGraw-Hill, 1968.
- Grimson, E. and Marr, D. (1979) "A Computer Implementation of a Theory of Human Stereo Vision," *Proceedings of ARPA Image Understanding Workshop*, S.R.I., pp. 41-45.
- Hildreth, E. (1980) "Implementation of a Theory of Edge Detection," *M.I.T. MS. Thesis*, Feb. 1980.
- Hubel, D and Wiesel, T. (1962) "Receptive Fields, Binocular Interaction and Functional Architecture of Monkey Striate Cortex," *J. Physiol.*, Vol. 195, pp. 215, 243.
- Julesz, B. (1971) *Foundations of Cyclopean Perception*, The University of Chicago Press, 1971.

Le Gros Clark, W. E. (1941) "The Laminar Organization and Cell Content of the Lateral Geniculate Body in the Monkey," *J. Anat.*, Vol. 75, pp. 419—433.

Le Gros Clark, W. E. (1942) "The Cells of Meynert in the Visual Cortex of the Monkey," *J. Anat.*, Vol. 76, pp. 369—377.

Logan Jr., B. F. (1977) "Information in the Zero-crossings of Bandpass Signals," *The Bell System Technical Journal*, Vol. 56, pp. 487—510.

Marr, D. (1976) "Artificial Intelligence - A Personal View," *M.I.T. A.I. Memo No. 355*.

Marr, D. (1977) "Representing Visual Knowledge," *M.I.T. A.I. Memo No. 415*.

Marr, D. and Hildreth, E. (1979) "Theory of Edge Detection," *M.I.T. A.I. Memo No. 518*.

Marr, D. and Nishihara, H. K. (1978a) "Visual Information Processing: Artificial Intelligence and the Sensorium of Sight," *Technology Review*, Vol. 81, No. 1, pp. 1—23.

Marr, D. and Nishihara, H. K. (1978b) "Representation and Recognition of the Spatial Organization of Three—Dimensional Shapes," *Proc. R. Soc. Lond.*, Vol. 200, pp. 269—294.

Marr, D. and Poggio, T. (1976) "From Understanding Computation to Understanding Neural Circuitry," *M.I.T. A.I. Memo No. 357*.

Marr, D. and Poggio, T. (1977) "A Theory of Human Stereo Vision," *M.I.T. A.I. Memo No. 451*.

Marr, D. and Poggio, T. (1979) "A Computational Theory of Human Stereo Vision," *Proc. R. Soc. Lond.*, Vol. 204, pp. 301—328.

Marr, D., Hildreth, E. and Poggio, T. (1980) "The Smallest Channel in Early Human Vision," *J. Opt. Soc. Am.*, (in press).

Marr, D. and Ullman, S. (1979) "Directional Selectivity and its use in Early Visual Processing," *M.I.T. A.I. Memo No. 524*.

- Poggio, T. (1980) "Trigger Features or Fourier Analysis in Early Vision: A New Point of View," in *The Role of Feature Detectors*, Gough and Peters eds., (in press).
- Sutherland, N. S. (1979) "The Representation of Three-Dimensional Objects," *Nature*, Vol. 278, pp. 395—398.
- Ullman, S. (1979) "The Interpretation of Structure from Motion," *Proc. R. Soc. Lond.*, Vol. 203, pp. 405—426.
- Westheimer, G. (1976) "Diffraction Theory and Visual Hyperacuity," *Am. J. Optometry Physiological Optics*, Vol. 53, No. 7, pp. 362—364.
- Westheimer, G. (1977) "Spatial Frequency and Light-Spread Descriptions of Visual Acuity and Hyperacuity" *J. Opt. Soc. Am.*, Vol. 67, No. 2, pp. 207—212.
- Westheimer, G. and McKee, S. P. (1975) "Visual Acuity in the Presence of Retinal-Image Motion," *J. Opt. Soc. Am.*, Vol. 65, No. 7, pp. 847—850.
- Westheimer, G. and McKee, S. P. (1977a) "Integration Regions for Visual Hyperacuity," *Vision Research*, Vol. 17, pp. 89—93.
- Westheimer, G. and McKee, S. P. (1977b) "Spatial Configurations for Visual Hyperacuity," *Vision Research*, Vol. 17, pp. 941—947.
- Westheimer, G. and McKee, S. P. (1978) "Stereoscopic Acuity for Moving Retinal Images," *J. Opt. Soc. Am.*, Vol. 68, pp. 450—455.
- Wilson, H. and Giese, S. C. (1977) "Threshold Visibility of Frequency Gradient Patterns," *Vision Research*, Vol. 17, pp. 1177—1190.
- Zeki, S. M. (1974) "Cells Responding to Changing Image Size and Disparity in the Cortex of the Rhesus Monkey," *J. Physiol.*, Vol. 242, pp. 827—841.

